SPEECH MONOGRAPHS

Published by THE SPEECH COMMUNICATION ASSOCIATION

VOLUME XXXVIII NOVEMBER, 1971 No. 4

THE DEVELOPMENT OF AN INSTRUMENT FOR MEASURING INTERACTION BEHAVIOR IN SMALL GROUPS

JAMES C. McCROSKEY and DAVID W. WRIGHT

ALTHOUGH many writers have suggested that more research in communication should consider process, relatively few reported studies have actually measured communication behavior from a process orientation. This is particularly true of the research on small group communication. A partial explanation for this shortcoming in small group research may be that researchers have relied too extensively on Bales' Interaction Process Analysis (IPA).[1]

IPA has several major limitations. Gouran, for example, has pointed to two major weaknesses: First, Bales' categories are mutually exclusive, a characteristic of the system which prevents a contribution from being classified in more than one way, and which presumes unidimensionality of individual contributions. Second, the system yields data that cannot be subjected to normal parametric statistical analysis.[2]

Leathers recently reported a new instrument which he has suggested as "an alternative to product measurement by attempting to measure the immediate effect of different types of contributions on group communication."[3] Leathers' approach holds much more promise for measuring communication behavior from a process view than IPA because it has the potential for overcoming both of the major weaknesses of the IPA system noted above. The Leathers instrument permits raters to respond to individual interaction behaviors of communicators in a small group on nine "dimensions," using semantic differential-type scales.

The major problem with the Leathers instrument is that, although he was able to obtain fairly high reliability from his raters, the "dimensions" of the instrument were apparently subjectively determined. No data from factor analysis has been reported to support the existence of these supposed "dimensions."

The purpose of the present study was the development and testing of an instrument for measuring interaction behavior in small group communication. It was suspected at the outset that such behaviors are multi-dimensional in nature. This study, therefore, was designed to uncover those dimensions and produce a reliable and valid measure for each dimension, one that would yield data amenable to parametric statistical analysis.

METHOD

Materials A thirty-minute discussion on the topic "What should the univer-

Mr. McCroskey is Associate Professor and Director of Graduate Studies in Communication; Mr. Wright is Assistant Professor of Communication, Illinois State University.

1 Robert F. Bales, *Interaction Process Analysis, A Method for the Study of Small Groups* (Reading, Mass., 1950).

2 Dennis Gouran, "Conceptual and Methodological Approaches to the Study of Leadership," *Central States Speech Journal*, XXI (1970), 222.

3 Dale G. Leathers, "Process Disruption and Measurement in Small Group Communication," *QJS*, LV (1969), 287.

sity do about parking in the campus area?" was videotaped. The discussants were five undergraduate students in their first course in small group communication at Illinois State University. Participation was spontaneous and unrehearsed. The discussants were aware that they were being videotaped, but reported that they did not feel that the videotaping interfered with their normal communication patterns.

A thirty-item, seven-step semantic differential-type instrument was developed. The instrument included the nine items previously employed by Leathers.[4] Two items similar to each of the Leathers items were included. These were added to increase the possibility that each of the Leathers items, in company with the added items, could generate an independent factor from factor analysis, if indeed such a dimension existed. Three general items were also added. The items on the instrument were as follows (the Leathers items are capitalized):

Wordy:short, INFLEXIBLE:FLEXIBLE, uncritical:critical, obstructive:constructive, SIGNAL:SYMBOL, fragmented:whole, task oriented:socially-emotionally oriented, IDEATIONAL:PERSONAL, INVOLVED:WITHDRAWN, feeling response:thinking response, interested:apathetic, logical:non-logical, tangential:goal-bound, bothered:cool, uncompromising:compromising, complete:incomplete, brief:lengthy, disorganized:organized, CLEAR:CONFUSED, unconcerned:concerned, harmful:helpful, related:unrelated, DIGRESSIVE:CONCISE, RELEVANT:IRRELEVANT, concrete:abstract, uptight:calm, unchangeable:changeable, ATOMIZED:UNIFIED, TENSE:RELAXED, ill-defined:well-defined

Evaluators There were two phases to the current investigation, the initial phase and the replication. In the initial phase two groups of evaluators were employed. Each group was composed of thirteen students in their first undergraduate course in small group communication. In the replication twelve students in an advanced graduate seminar in small group communication were employed as evaluators. The evaluators received brief training in the use of the instrument before being employed in the study. There was general discussion of the meaning of the terms used on the instrument. A taped discussion, similar to the videotape used in this study, was played so that the evaluators would have practice in use of the instrument. Procedural and semantic problems were covered, as well as other administrative details.

Procedures Fifteen stimulus statements were randomly selected from the videotaped discussion for the initial phase of the study. A table of random numbers and the tape-distance counter on the video recorder were employed in the selection. The evaluators viewed the videotape until it was stopped immediately after a stimulus statement. Their attention was called to the stimulus statement by rewinding the videotape and replaying the statement. They were asked to complete the evaluation instrument on the basis of the next participation following the stimulus statement (the response).

Two minutes were allotted for the evaluators to complete the instrument after each stimulus-response induction. In every instance the evaluators had ample time to complete the instrument. The total administration time for each group of evaluators was approximately one hour.

The same procedures were followed for both the initial phase of the study and the replication, except that different stimulus statements were selected for the two phases.

[4] Leathers, pp. 287-300.

Statistical Analysis

The data obtained from the two phases of the study were analyzed separately. In each case the data were first submitted to principal component factor analysis and varimax rotation. The cut-off criterion for rotation was an eigenvalue of 1.0. An item was considered loaded on a given factor if it had a rotated factor loading on that factor of at least .60 and had no rotated loading on another factor higher than .40.

After the factor structure had been determined, the two items with the highest and purest loadings on each of the factors (based on the above criteria) were selected and scored for each evaluator on each of his fifteen responses. The reliability of the evaluators' use of the instrument was examined by means of the analysis of variance procedure proposed by Hoyt and Guilford.[5]

Finally, the data were subjected to one-way analysis of variance with repeated measures, the fifteen stimulus-response evaluation points serving as the levels of the independent variable. The data for each factor were analyzed separately.

With twenty-six evaluators each completing the instrument fifteen times, the "N" for the initial phase of the study was 390. The "N" for the replication was 180, twelve raters completing the instrument fifteen times.

It should be noted that fifteen completed measures were entered into the factor analysis for each rater. This procedure confounded the between and within subjects variances. An alternative method of analysis which avoids this problem has been suggested by Tucker[6] and employed with semantic differential data by Reid.[7] This procedure not only determines the essential factor structure but also helps determine concept groupings and factors most closely associated with each of those concept groupings. While this approach would normally be preferred to the procedure employed in the present study, it was not employed because of the unusual nature of the "concepts" in this study. The "concepts" were fifteen randomly selected interactions. Since any grouping of these interactions would have been difficult, if not impossible, to interpret, the regular principal, component factor analysis procedure was employed.

As a result of the procedure selected it may be argued that the resulting factor structure may not apply to all types of interaction. While this indeed may be true, subanalyses of the present data did not produce results suggesting this conclusion. Separate factor analyses for each of the fifteen stopping points in the initial phase of the study were computed. While there were insufficient data ($n = 26$) to insure stable factor analyses, in each case the factor structure produced was essentially the same as that for the combined analysis. Similarly, the factor structures obtained from the initial phase of this study and the replication were almost identical, even though fifteen different interactions were rated in the replication.

Results

Factor analysis indicated the presence of six factors on the instrument in both the initial phase of the study and in the replication. These factors were labeled,

[5] J. P. Guilford, *Psychometric Methods*, 2nd ed. (New York, 1954).

[6] Ledyard R. Tucker, "Some Mathematical Notes on Three-Mode Factor Analysis," *Psychometrika*, XXXI (1966), 279-311. See also "Three-Mode Factor Analysis of Parker-Fleishman Complex Tracking Behavior Data," *Multivariate Behavioral Research*, II (1967), 139-51.

[7] J. Christopher Reid, "A Three Mode Factor Analysis of Students' Perceptions of a University," *Journal of Experimental Education*, XXXVIII (1969), 93-96.

on the basis of the content of the items with the highest rotated loadings on the factors, as follows: orientation, tension, flexibility, relevance, interest, and verbosity. Table 1 reports the items selected for the Interaction Behavior Measure (IBM) and the rotated factor loadings for each item on each factor for both the initial study and the replication.

Analysis of variance reliability estimates were computed separately for the twenty-six raters in the initial phase of the study and for the twelve raters in the replication. The obtained reliability estimates for each factor composed of the two best items on that factor are reported in Table 2. The obtained reliability estimates ranged from .64 to .92.

The factor analytic procedure is designed to discover independent dimensions present in an instrument such as the one employed in this investigation. In theory at least, such dimensions should be uncorrelated. To the extent that they are correlated, their potential usefulness is reduced. In order to determine how independent the factors obtained in this study were, correlations

TABLE 1
INTERACTION BEHAVIOR MEASURE

Item	Evaluation Group	Rotated Factor Loadings** Orientation	Tension	Flexibility	Relevance	Interest	Verbosity
Task Oriented:	Initial	.81*	.06	.01	.31	.17	.00
Socially-Emotionally Oriented	Replication	.61*	.00	.06	.39	.22	.11
Ideational:Personal	I	.81*	.03	.10	.19	.01	.17
	R	.78*	.01	.05	.22	.17	.16
Bothered:Cool	I	.08	.75*	.09	.16	.29	.20
	R	.04	.87*	.04	.20	.04	.02
Tense:Relaxed	I	.02	.75*	.18	.24	.18	.14
	R	.09	.82*	.09	.21	.05	.23
Flexible:Inflexible	I	.02	.06	.75*	.07	.21	.15
	R	.22	.00	.79*	.02	.01	.09
Unchangeable:Changeable	I	.07	.12	.67*	.17	.18	.13
	R	.05	.11	.87*	.04	.00	.09
Relevant:Irrelevant	I	.26	.07	.11	.73*	.34	.03
	R	.16	.06	.08	.84*	.25	.09
Related:Unrelated	I	.14	.01	.12	.76*	.39	.07
	R	.26	.15	.05	.76*	.30	.11
Interested:Apathetic	I	.15	.20	.02	.17	.75*	.07
	R	.26	.20	.05	.30	.73*	.06
Involved:Withdrawn	I	.08	.20	.02	.24	.76*	.04
	R	.34	.02	.03	.26	.71*	.00
Wordy:Short	I	.04	.19	.04	.23	.01	.83*
	R	.10	.13	.01	.00	.09	.91*
Brief:Lengthy	I	.04	.20	.03	.20	.07	.88*
	R	.15	.18	.02	.03	.04	.89*

*Highest Loading
**Rounded to two places, sign ignored.

TABLE 2
EVALUATOR RELIABILITY ESTIMATES

Group	Factor: Orientation	Tension	Flexibility	Relevance	Interest	Verbosity
Initial (n = 26)	.92	.66	.64	.71	.86	.78
Replication (n = 12)	.88	.87	.69	.74	.68	.74

TABLE 3
INTERFACTOR CORRELATION MATRIX

Group	Factor	Factor				
		1	2	3	4	5
Initial (n = 390)	2	—.16				
	3	—.03	.08			
	4	.56*	—.34*	—.09		
	5	.52*	—.25*	.01	.58*	
	6	—.06	—.07	—.05	—.04	.06
Replication (n = 180)	2	—.14				
	3	—.06	.09			
	4	.46*	—.31*	—.05		
	5	.29*	—.41*	—.02	.50*	
	6	—.01	—.02	.02	—.02	.05

*Correlation significant at .05 level.

among the scores on the various factors for each of the two phases of the study were computed. The results of this analysis are reported in Table 3. While factors 1, 4, and 5 were significantly intercorrelated, as were factors 2, 4, and 5, the maximum variance on one factor predictable from another was only approximately thirty per cent.

The contributions randomly selected for evaluation in this investigation differed markedly in type and quality. If the IBM is to be presumed to have any validity for the description of interaction behavior, it should reflect those differences. The repeated measures analyses of variance provided a direct test of the hypothesis that the interaction behaviors evaluated in this study differed on the six factors of the IBM. The results of these analyses provided support for that hypothesis for all of the factors except verbosity. With the exception of this factor, all of the analyses yielded F-ratios that were statistically significant at the .01 level. The results on the verbosity factor were clearly not significant ($F < 1.0$).

DISCUSSION

The first purpose of this investigation was to discover the dimensions of interaction behavior in small group communication. Six dimensions were discovered in the initial investigation and the same six dimensions were observed in the replication. Because the subjects in the initial investigation were comparatively untrained in small group communication theory (only part way through their first course) while those in the replication were highly trained (in their second graduate course), it is reasonable to conclude that the observed dimensions are not the function of instruction in small group communication theory. The IBM, therefore, can be used by evaluators with either minimal or extensive knowledge of small group communication theory with the expectation that the factor structure in the resulting data will be essentially the same. This is particularly important for the researcher who has few or no potential evaluators available who are knowledgeable in small group communication theory.

The second purpose of the study was the development of a measuring instrument for interaction behavior in small group communication that would be amenable to parametric statistical analysis. The semantic differential approach to measurement has been generally accepted by researchers as one which yields data that meet the assumptions re-

quired for parametric statistical analysis; hence, this approach was selected.

The resulting instrument was found to have acceptable reliability on each of the six factors measured. Since reliability in the use of this type of instrument is closely tied to the number of evaluators using the instrument, the researcher who wishes to improve his reliability of measurement may do so by increasing the number of evaluators he employs. Since people need no special background to be selected as evaluators, and training of evaluators is a brief and simple task, this is a viable procedure.

Validity is always an important question in the development of a measuring instrument. In most cases, there is no absolute criterion against which to compare a newly developed instrument to test its validity. This instance is no exception. One important check on the validity of a measure is whether it can detect differences we know or believe exist. Differences were detected among the fifteen interaction behaviors employed in this study on five of the six dimensions of the instrument. The exception was the verbosity dimension. Since this dimension is primarily concerned with the length of a contribution, if all of the contributions were of about the same length, no difference in verbosity scores should be expected. This was the case in the present study.

The results of this study, therefore, indicate that there are six observable dimensions of interaction behavior in small group communication and that the IBM is capable of reliably measuring those dimensions with some presumption of validity. The IBM permits the researcher to examine the effects of any number of communication variables in small group communication from a process orientation in a way that yields data that can be analyzed by means of any appropriate parametric statistical procedure.

Although the IBM was developed primarily as a measure of interaction behavior in intragroup communication, it should be equally applicable to measurement of such behavior in any interpersonal communication setting. One limitation, however, must be stressed. The IBM was developed in a setting where evaluators could respond to both verbal and nonverbal stimuli provided by communicators. Therefore, the instrument should be useful for measuring interaction behaviors observed live or on videotape. The use of audiotape, however, would remove many of the nonverbal stimuli, and manuscripting would remove even more. Hence, application of the IBM in these circumstances must be done with full awareness that some of the dimensions (particularly tension and interest) may not be measured as well as they would be if the live or videotape approach were employed.